DOCUMENT RESUME

ED 231 370 IR 050 244

AUTHOR Pejtersen, Annelise Mark
TITLE Design of a Classification Scheme for Fiction Based on an Analysis of Actual User-Librarian Communication, and Use of the Scheme for Control of Librarian's Search Strategies.
PUB DATE Aug 77
NOTE 22p.; Paper presented at the International Research Forum on Information Science (2nd, Oslo, Norway, August 1977). For related document, see IR 050 245.
PUB TYPE reference Materials - Vocabularies/Classifications/Dictionaries (134) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Check Lists; *Classification; Content Analysis; *Fiction; Foreign Countries; Library Research; Library Services; Literary Genres; *Public Libraries; *Reading Material Selection; Search Strategies
IDENTIFIERS Denmark; *User Needs

ABSTRACT

Because the general alphabetic arrangement of fiction in libraries is of little use to the ordinary reader, whose criteria for selecting books often relates to book content, a multi-dimensional and inexclusive fiction classification scheme was developed as a model for both book description and search strategy. Based on analysis of 160 actual user-librarian conversations about fiction recorded under everyday library conditions in Danish public libraries in 1973-74, and a further analysis of 134 conversations recorded in 1976, it was found that users' subconscious classification of fiction can be characterized by means of four dominant dimensions: (1) subject matter, including the categories of action and the course of events, psychological development and description, and social relations; (2) type of frame, including time and geographical/social environment frames; (3) author's intention, including provision of an emotional experience or provision of information; and (4) accessibility, which includes readability and physical characteristics of a book. This report describes the development and applications of the prototype classification scheme based on these four dimensions. Appendices present the classification scheme in detail with examples of its classes and subclasses. Also shown are examples of user formulations and genre-terms to illustrate how these criteria are congruent with the various divisions of the scheme. (Author/ESR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☞ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED231370

DESIGN OF A CLASSIFICATION SCHEME FOR FICTION BASED ON AN ANALYSIS OF ACTUAL USER-LIBRARIAN COMMUNICATION, AND USE OF THE SCHEME FOR CONTROL OF LIBRARIAN'S SEARCH STRATEGIES

Annelise Mark Pejtersen

IR050244

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Annelise Mark Pejtersen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Design of a classification scheme for fiction.

1. Introduction: The problem

The classification of fiction in public libraries is based on an alphabetical arrangement by author supported by a card catalogue together with a cross-reference card catalogue after titles. This offers the advantage of uniqueness in the placing and finding of an individual book in the library, but the alphabetical arrangement of fiction is of little help to the librarian, when it comes to serving the users, whose needs typically are centered around other aspects of book and author than the alphabetical location of the writer. These aspects manifest themselves in the users' different questions which are not formulated according to any well-known systematism, but only as a wish to get, for example, "an exciting book, but not with too many problems or politics". There is no exact or unique answer to this question, nor is there within the present library system auxiliary tools, which the librarian can use to establish a base for his search strategy. Furthermore the librarian's analysis of the user's needs and his corresponding search strategy are complicated by the user's ambitions to keep up with the library system, as he imagines it. Users formalize their needs like this: "a small book - with big letters, placed under h" or they search for books after primitive criteria like "red books" or "worn out books". This illustrates how strange the library system is to the user.

In the communication with the user the librarian must operate with two different schemes of classification. The alphabetical classification of the library, which is useful in the retrieval of specific books by author and title, and the subconscious and intuitive classification of the users from different points of view based on their needs and interests. The librarian faces two main difficulties; first the problem of identifying the user's needs, and then the problem of formulating a relevant search strategy among books which are not classified according to needs, but alphabetically according to the author.

In the librarian's search strategy and his communication with the user, the genre-classification generally serves as a frame of

reference, and consequently some libraries have set up a genre-catalogue. The traditional genre-classification characterizes only one aspect of the book (example: love-story) and is insufficient in front of the multi-dimensional needs of the users. It is inadequate because it is an exclusive classification, which ignores relevant combinatorial possibilities of fitting books to users' needs. The consequence is that a book is either a crime novel or a historical novel, although in fact it may be both. The employment of heterogeneous criteria in the genre-classification as seen from designations like love-story, (a subject-matter criterion) historical novels, (a criterion of time) thrillers, (a criterion of emotion) social criticism, (a criterion of author's intention) already indicates several dimensions relevant to the ordinary reader. (Figure 1). The exclusive and heterogeneous genre-classification has functional implications, which makes it ineffective as a model for search strategy and communication.

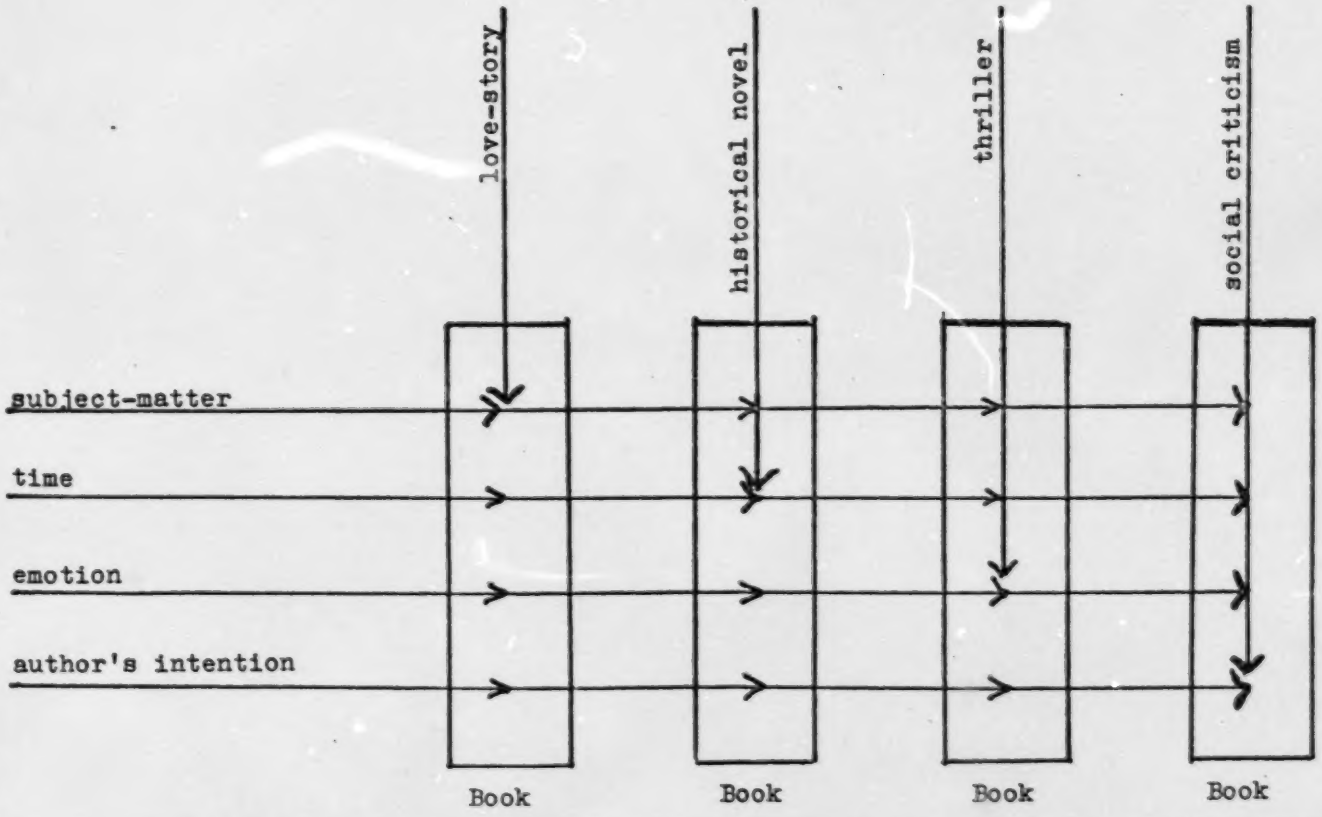
It is one of the most important tasks of the librarian to create a system of reference or a translation between the user's classification and that of the library, i.e.: to take the users' information about certain characteristics of a book and translate this information into a useful entry in the library system. This, however, poses several problems.

1. The users' classification is unsystematic, so the librarian will have to create for himself an adequate and superordinate system which is compatible with users' classification. In this way it becomes possible to ask the user systematically for further information which finally leeds to an identicifation of the users' request within the librarian's frame of reference.
2. The main part of the request/book correlation is usually rather intuitive and it is most likely that many of the habits of the user and the librarian are based on subconscious attitudes and associations which cannot be consciously formulated.

This problem can probably be solved by a classification based on a formulation of several of the aspects, which characterize the books in relation to users' needs. Such a classification should define the boundaries and attributes of a reasonable and handy number of classes within these aspects.

If a classification system is used which does not correspond to

Figure 1.



The heterogeneous and exclusive genre-classification.

the users' unconscious classification, and the librarian is asked to find a book with such and such characteristics, he will face the problem that the request will have to be redefined and translated into his classification system.

This translation may be avoided if we can design the classification of the librarian from an analysis of the classification demonstrated by users in real library situations. This will make it possible for the librarian to step more directly from the user's request to the target material, perhaps after a few additional questions in order to clarify formulation and meaning, or to get information in more dimensions.

As the system of reference must be operational i.e. it must aim at a realistic procedure in serving the user, studies of user-librarian interactions have been necessary for a realistic development of a user-oriented classification scheme.

2. Investigation of users' classification of fiction.

The following conclusions are extracts from a study of 160 actual user-librarian conversations about fiction recorded under everyday library conditions in Danish public libraries in 1973-74 (ref. 3). The results of the analyses have been confirmed by another investigation comprising 134 conversations recorded in 1976 (ref. 4). In both cases the user-librarian conversations were written down literally by students accompanying a librarian who was aware of the experiment. The user on the contrary was ignorant of the student's presence and usually did not notice it.

The analysis of frequency is users' formulations and needs.

The subconscious classification of fiction made by almost three hundred users can be characterized by means of four dominant dimensions (Figure 2):

Figure 2.

Users' formulations

Frequency per 100 conversations.

Dimensions	Frequency	Examples
1. subject-matter	38	books with psychology
2. frame time/place (geographical, social environment, professions)	25	books about old days, books about castles and mansions
3. author's intention/ attitude	37	romantic books, relaxing and easy books, love story, preferably philosophical
4. Accessibility	34	easy books, not too heavy books with big letters, not with complicated names
other formulations	87	could you advise me in reading modern fiction? a book by Max Frisch or something similar.

The four dominant dimensions of users' subconscious classification of fiction.

The users' choice of words has been translated literally.

The category "other formulations" includes all expressions which are not included in the 4 dimensions, and represent terms by which the users most frequently characterize books. (Examples from the category of "other formulations": "I have read Russian literature before", specific titles/authors, "the same genre", "good books", etc.) An analysis of users' expressions classed as "other formulations" reduces this category considerably, since the term "good book" covers an interest in family chronicles, for instance.

The dominant dimensions: Subject-matter, frame, author's intention and accessibility can be further specified on the base of the users' different expressions and needs as seen in figure 3.

When contacting the librarian, the users either describe their wishes directly within the categories shown in figure 3, or the subsequent conversation with the librarian reveals implicitly their actual desires in terms of these categories. All formulations of the users are represented in this scheme, except the question for "modern literature/new literature", unless we establish a subclass called 'first edition' under physical characteristics. Though, in many cases these phrases conceal a need for certain qualities identical with the other dimensions in figure 3.

The four aspects or points of view do not exist as four isolated independent needs at the same level and with the same importance in the users' classification - even if requests based on these dimensions are likely to be formulated as being of equal importance. Generally speaking, a user will expect to have his reading need satisfied by an emotional experience, for instance by an exciting or sentimental book, or he will expect his reading need to be satisfied by cognition and information, for instance about human and social conditions. The users' expectations of an emotional experience and cognition/information are correlated with the subject-matter and the frame as well as with the author's attitude to the subject dealt with in the book. A further decisive factor for a successful reading experience is the accessibility of the book, with regard to the level of communication employed by the author.

A user's change of need from an emotional experience to cognition/information will often cause a change in his demand on the subject-matter and frame of the book. On the other hand there seems to be

Figure 3.

Users' formulations

Frequency per 100 conversations

Dimensions	frequency	subclasses	frequency	Examples
1. subject matter	38	Action and course of events ("plot")	10	mystery novel, book with action
		psychological development/description	5	love-story, book with psychology
		social relations	19	family chronicle, not with politics
2. frame	25	Time	13	historical novel, books from 1600-1700 century
		Place (geographical, social environment, professions)	12	traveloque, books from the countryside, books about working people
3. Author's intention/ attitude	37	emotional experience	34	humoristic, suspense, amusing
		cognition/information	3	philosophical, political, not too many problems
4. Accessibility	34	readability	16	easy, not complicated, not heavy
		physical characteristics	18	typography, modern/old, series, size, volume
Other formulations	87	author's name/title	25	
		author's name/title as example	22	something like "Emily Brontë"
		good book	32	
		diverse	8	

The frequency of the users' formulations in the subclasses of the four dominant dimensions. The users' choice

no correlation between the user's expectations of emotional experience and cognition/information and the accessibility of the book sought. A change in the users' needs implies no change in the level of acquisition. Accessibility is not an end, but a means for the user.

The majority of users' formulations along the dimension of author's intention belongs to the class: emotional experience, which presumably is due to a difficulty in formulating needs about more abstract qualities like ideology and philosophy in books. Instead the user probably utilizes earlier experiences with correlations between emotional experiences/subjects and author's attitude to and intention with the subject.

A request for a family chronicle is often correlated with a wish for an author's intention related to cognition/information confirming the traditional and established values of society and family life.

A wish for "easy and entertaining books" informs the librarian indirectly about the user's lack of interest in books dominated by intentions of cognition and information. But an example from the user of "an easy and entertaining" book will reveal to the librarian what kind of author's attitude the user prefers.

Since all four aspects are of importance to the user - at different levels - a classification scheme must comprise all the relevant dimensions.

A curious example is the user who addresses himself to the librarian with a press cutting asking the librarian to find the book mentioned in the cutting, which is reproduced here:

Men vil man forstå mennesker, er det ikke nok, at betragte dem og tale med dem. Man må også læse om dem, fordi digterne kan hjælpe os til en dybere forståelse, end man selv ved egen hjælp kan finde frem til.

For nylig læste jeg en bog af en russer med et svært navn om mennesker på et kræfthospital.

Jeg troede ikke, at den ville interessere mig og var på forhånd bange for, at emnet var for dystert og for trist, og så viste det sig, at det var den bedste

But in order [1]to understand human beings it is not enough to watch them and speak to them. We must also read about them, [1]because poets can help us to achieve a deeper understanding than we can by ourselves.

Recently I read a book by a Russian with a difficult name [2]about people in a [3]hospital for cancer patients. I did not believe that it would interest me and was afraid beforehand that the subject was too [4]sinister and sad, and then it turned out that it was the best...

English translation

This clipping is delivered in the library in 1973.
Source unknown.

This cutting is a kind of subject-order, but the user wants a precise author and title. In the cutting the book is characterized by the same dimensions as we have found in the user-librarian conversations:

1: author's intention : cognition/information
2: subject matter : social relations
3: frame, place : social setting
4: author's intention : emotional experience

This information has had the effect of an appetizer to the user and appears directly as the user's criterion of choice. Exactly this information has been cut out by the reader and presented the librarian as relevant to the search procedure.

3. A classification scheme for fiction

A classification scheme based on an analysis of the relations between the formulations and actual needs of readers of fiction in public libraries must satisfy the demands of users and be specified to a degree, which is relevant to the user. At the same time it must be restricted to a relatively few number of criteria of classification sufficient to cover the librarian's negotiation with the user.

The following dimensions and classes should fulfil these conditions

I. subject-matter
   a. action and course of events
   b. psychological development and description
   c. social relations

Subject-matter, defined as the matter dealt with in the book.

II. frame
   a. time: past, present, future
   b. place: geographical, social environment, profession

Frame defined as the frame in time or place/environment which the author has chosen for his work. It is "the setting" so to speak aound the subject which it supplements. At the same time it locates the subject within a certain time-environment relation.

III. Author's intention

a. emotional experience

b. cognition and information

Author's intention. Two characteristics seem to be of importance for users: emotional experience and cognition/information, i.e. the attitude of the writer towards the subject and the set of ideas and emotions, which the author intends to bring to his readers.

IV. Accessibility

a. readability

b. physical characteristics

Accessibility defined as the level of communication of the book judged by qualities that generally facilitate or inhibit communication.

The acquisition of a book seen from a user's point of view is influenced by two factors: Readability, a linguistic measurement and physical characteristics which refer to physical qualities of importance to the reading process, as for instance typography. The readability classification is made by already established methods for analysing linguistic factors of importance to the accessibility of a book, e.g. readability indexes.

4. Characteristic of the classification scheme

This classification scheme is not exclusive since all characteristics may be found in the same book but with different dominance. Consequently a book may be classed in several places in the system, depending on the weight placed upon the different aspects.

The advantage of a non exclusive classification appears when considering a book, which is for instance both a love story, a historical novel and an exciting and romantic story. In this case it is not necessary to rank, since all aspects must be classed.

All subclasses along the following dimensions: subject-matter - frame - author's intention are not exclusive, but the dimension of accessibility employs exclusive subclasses. A book is either easy or difficult as to language, a book is either a novel or a short story, etc.

There is no logical order in the established classes; one class is not logically dependent on the other classes. So there is no hierarch

of classes. The classes are coordinate and do not comprise uniform subjects nor do they belong to one definable group. This makes it easier to meet the user's needs which are not centered around one single aspect of the book. The one-sided stressing of the hierarchical relation found in the traditional classification is set aside so as to take the advantage of a combination of some of the functions and aspects of fiction. The librarian can enter the system anywhere without first having to place the subject in some superordinate class and then search on for a subordinate class. The classification system is characterized by the fact that the classes, which are established, are not locked in a hierarchy in the system. On the contrary, any relation between designations and subjects may be expressed. And any novel can be classed, not only light literature or traditional genres.

As already mentioned, the classification-system aims at being operational in daily routine; therefore the number of dimensions is limited to four. But even with a low number of subdivisions in the four dimensions, the resolution of the scheme will be high due to the number of possible combinations.

An example of classification:

Pierre Boulle: "La Planète des Singes"
French. 1963.

Classification:

I.	Subject-matter	
	a. action and course of events	fight and experience during a space trip
	b. social relations	Monkey society and its problems
II.	Frame	
	a. time	future
III.	Author's intention	
	a. emotional experience	exciting
	b. cognition/information	criticism of earthly social conditions
IV.	Accessibility	
	a. readability	easy
	b. physical characteristics	novel, small typography, 141 pages, translated, Spektrum pocket books. 1971. Science-fiction. First edition 1963.

5. Use of the classification scheme

Search strategy

The user formulates his needs form his experience with a narrow selection of books. His questions to the librarian are often of low precision when seen in relation to his real needs and the actual contents of the book stock. In addition to different social backgrounds, semantic problems will often influence the negotiation with the librarian. The study of user-librarian communication illustrates how identical formulations can conceal different needs.

A demand for "books with much action" can cover a wish for as different materials as family chronicles and detective stories. Due to the semantic problem, a request for a science-fiction novel may cover a need for:

1. a book about futural technology: subject-matter dimension
2. an exciting book: emotional experience dimension
3. a social criticism: cognition dimension
4. an easy book: accessibility dimension

The science-fiction class in this four-dimensional example is the class which is limited by the subclasses along the four dimensions - and by offering suggestions along these four axes, it is possible to identify which of the dimensions really represents the user's need - and the librarian may then be able to satisfy the user with a ghost-story (emotional experience dimension).

The appendix on page 17 illustrates that a semantic analysis of user-formulations can reveal, how identical expressions conceal varying interests within different dimensions represented in the classification scheme.

Interview-technique

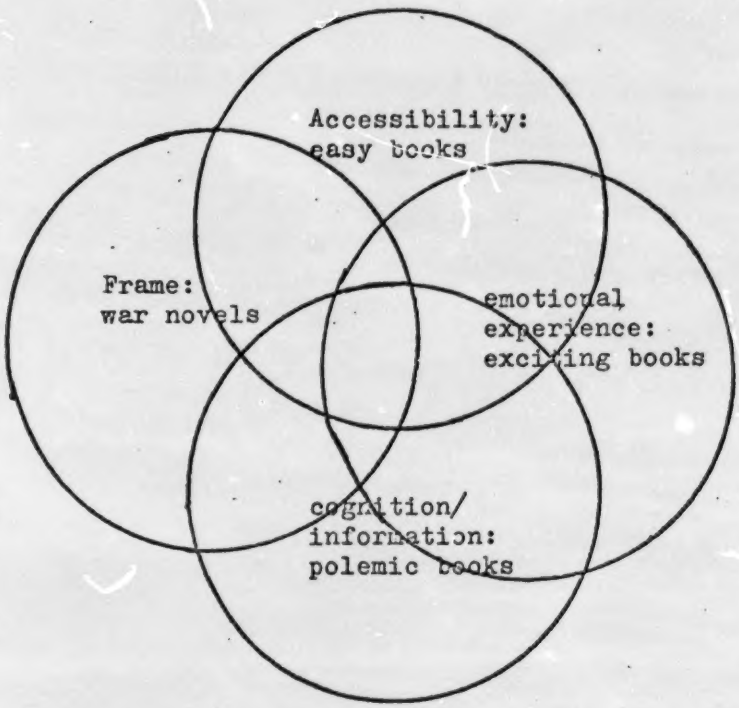
The efficiency of the classification scheme as a support to the librarian's search strategy is due to the fact mentioned previously that the classification is not exclusive.

This redundancy in the classification can be of great help to the librarian in his determination of a user's desire, since the title of the book may become some sort of cross reference initiating searching in fields not very obvious from the user's original request

If a user asks for a "war novel", the librarian, by means of additional questions, will have to make sure, whether the user means: a book about war, an exciting book, a polemic book or an easy book. The necessary supplementary questions must be asked. When the librarian has placed the information "war novel" in his reference system, he can ask relevant questions by searching dimensions and classes - possibly in the form of suggestions of books identified by the system. In the same way, the user's information about what he has read with interest can only be useful if it can be placed in the same frame of reference. Only after several questions and answers centered around a number of books or subjects with some common features, is the librarian able to conclude what the user's real needs are.

This technique of inquiry is necessary in order to ascertain, whether the user formulates his need in a broader sense than it actually is.

Figure 4



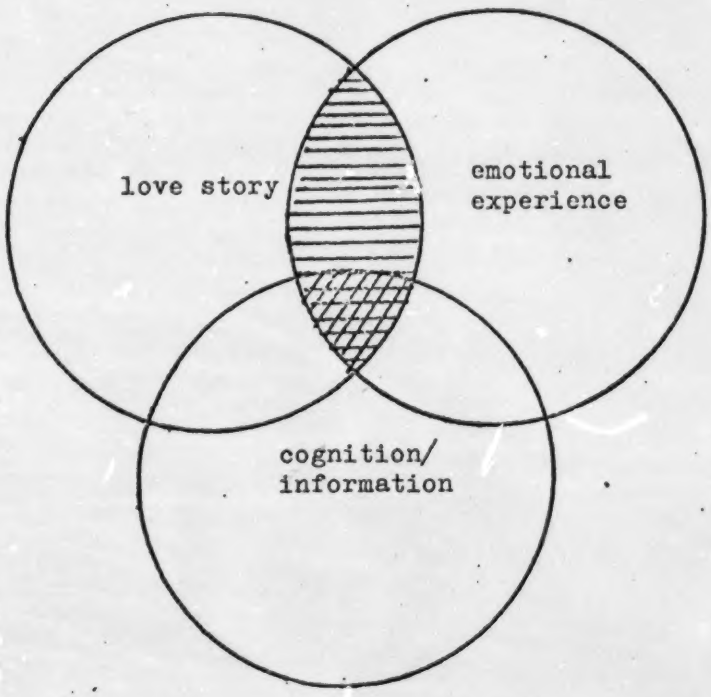
The figure demonstrates how the simultaneous determination of the reader's interest in several dimensions of a desired book can lead to the definition of very specific classes of books.

Cultural values

Apart from satisfying the users' precise and immediate need it is the aim of the librarian to communicate cultural values. The classification scheme can also be used to pursue this purpose, when users' questions are too specific in relation to their actual need, which might be more or less differentiated.

After a determination of a user's demand, for instance of two aspects, it is necessary to propose a book, which meets this demand in order to get the user's acceptance.

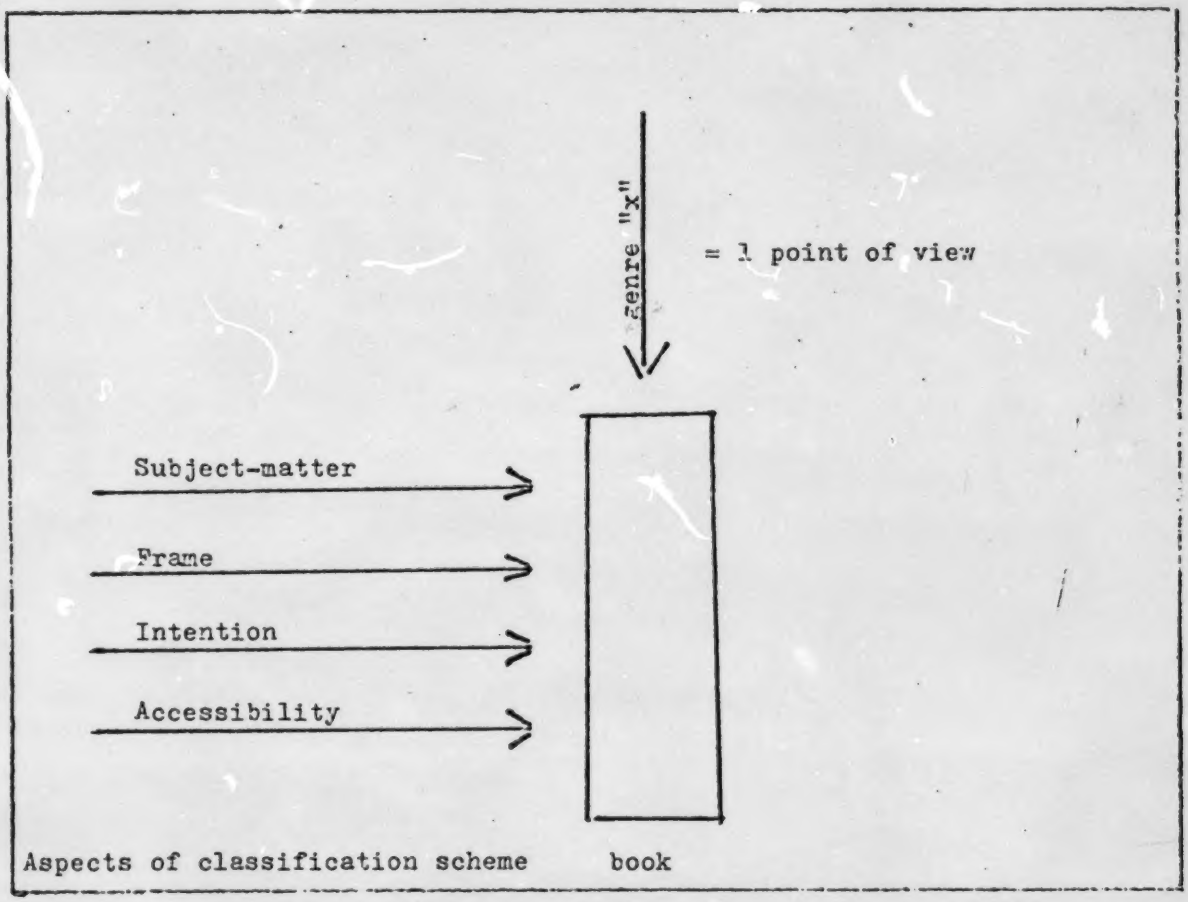
Figure 5



In this case the user desires a sentimental love-story, and the librarian can suggest a sentimental love-story, which also has an author's intention of cognition/information.

The librarian has the possibility and freedom to experiment within the other qualities of the book, which a multidimensional search gives information about, and thus help the reader to try new orbits and widen his reading interests. This cannot be obtained by the traditional genre-classification, which does not inform the librarian of more than one aspect of a book (figure 6).

Figure 6



The genre-classification inhibits the librarian in satisfying combinatorial user needs as well as the exchange of cultural values.

Conclusion

Keeping the dimensions of the classification scheme in mind when reading a book, supports the librarians' attention and entrances to a subsequent memory of user relevant aspects of the book. The system also makes it possible to extract this information by quick reading methods such as skimming. A systematic use of the classification scheme leads very easily to user-oriented catalogues or card-catalogues. Based on this scheme a card-catalogue is established in a public library, and a group of librarians are trying the function of a user-oriented catalogue on the same principles. Since 1973 the teaching of students comprises this classification model, and the development of heuristic rules to improve the search strategy and user-communication in fiction is discussed.

The problem of classifying fiction arose already in the 30ties. The idea of characterizing books through reading habits led J.H. Foster (ref. 1) to undertake a sociological investigation of different reading groups and the books, they preferred. However, the discussion has faded out and no serious effort to analyse users' needs and formulations in fiction has been published. Grolier states (ref. 2) in his review on classification in public libraries: "En outre, la classification doit, autant qu'il est possible, correspondre aux "motifs de lecture" des usagers, à leur attitudes vis-à-vis des livres." (p. 20).

The difficulty of finding an efficient scheme has probably caused a disinterest in the classification of fiction. This might explain the absence of discussion and proposals of methods to solve the problem.

Appendix

The classification scheme is shown in Figure 7-10. Each dimension consists of a set of classes or specifications, which possesses subclasses, examples of which are shown. The column of examples of user formulations and genre-terms illustrates how the criteria used in the traditional genre-terms are congruent with the various divisions of the scheme. (Example: historical novel is placed in the sub-class 'past' under the division of 'Frame': time). Simultaneously genre-terms are also placed in accordance with the semantic contents, which the terms sometimes express, when employed by users in the concersation with the librarian. (Example: Love-story is placed in the subclass 'exciting books' under the division of 'Author's intention': emotional experience.)

Figure 7

1. Subject-matter	
a. Action and course of events	Examples of user formulations and genre-terms
ex: expeditions	crime novel
crime - detection	historical novel
fight	war novel
invention, discovery	exciting book
sex	family chronicle
technical progress	love story
	travelogue
b. psychological development and description	
ex: childhood and adolescence	social novel
development	philosophical love story
Bildung	mystery novel
religious conflicts	sentimental book
	suspense novel
c. social relations	
ex: groups	science-fiction novel
family	religious novel
society	easy and entertaining book
	family chronicle

Figure 8

Frame	Examples of users' formulations and genre-terms
a. time	psychological novel
ex: past	historical novel
present	thriller
future	biographies
	science fiction
	family chronicle
b. geographical/social environment	
ex: place	western
profession	humorous novel
organisation	travelogue
	love story
	historical novel
	philosophical novel
	medical novel

Figure 9

Author's intention	Examples of users' formulations and genre-terms
a. emotional experience	
ex: horror/thrill	love story
amusement	thriller
relaxation	sea-novel
suspense	travelogue
romance	family chronicle
entertainment	suspense novel
	biography
	science fiction novel
b. cognition and information	
ex: description	social realistic novel
agitation	love story
authenticity	historical novel
	science fiction novel
	humorous novel
	war novel
	entertaining books
	spy story
	romantic novels

Figure 7-9. The traditional genre-terms and users' formulations seen in relation to the four dimensions and their subclasses.

Figure 10

Accessibility	Examples of user' formulations
a. readability	
ex: easy	easy to read
medium	short stories in weeklies
difficult	pulp magazines
composition	entertaining books
language	not too heavy books
b. physical characteristics	
typography	short stories
length/volume	big letters
novel	pulp magazines
short story	new literature
type of edition/time/place	series
first edition	
nationality	
editor	
anthology, series	

The common genre-terms cover the whole spectrum of readability, and normally have individual characteristics within one or more of the classes under physical characteristics.

Short stories in weeklies, pulp magazines etc. are not to be found in the library but are included here because the user often refers to them, thus revealing his literary level to the librarian.

Under physical characteristics the classification is based on the physical qualities that affect the accessibility of the work (magna print, etc.).

This includes also the type of edition which does not actually affect accessibility but is directly combined with it (e.g. pulp magazines etc.) and which gives the librarian some indication of the accessibility level.